Virus Genes (2017) 53:856–867
DOI 10.1007/s11262-017-1485-4

Characterization of *Eptesipoxvirus*, a novel poxvirus from a microchiropteran bat

Shin-Lin Tu[1] · Yoshinori Nakazawa[2] · Jinxin Gao[2] · Kimberly Wilkins[2] · Nadia Gallardo-Romero[2] · Yu Li[2] · Ginny L. Emerson[2] · Darin S. Carroll[2] · Chris Upton[1]

Received: 4 April 2017 / Accepted: 30 June 2017 / Published online: 6 July 2017

Abstract The genome of *Eptesipoxvirus* (EPTV) is the first poxvirus genome isolated from a microbat. The 176,688 nt sequence, which is believed to encompass the complete coding region of the virus, is 67% A+T and is predicted to encode 191 genes. 11 of these genes have no counterpart in GenBank and are therefore unique to EPTV. The presence of a distantly related ortholog of *Vaccinia virus* F5L in EPTV uncovered a link with fragmented F5L orthologs in *Molluscum contagiosum* virus/squirrelpox and clade II viruses. Consistent with the unique position of EPTV approximately mid-point between the orthopoxviruses and the clade II viruses, EPTV has 11 genes that are specific to the orthopoxviruses and 13 genes that are typical, if not exclusive, to the clade II poxviruses. This mosaic nature of EPTV blurs the distinction between the old description of the orthopoxvirus and clade II groups. Genome annotation and characterization failed to find any common virulence genes shared with the other poxvirus isolated from bat (pteropoxvirus); however, EPTV encodes 3 genes that may have been transferred to or from deerpox and squirrelpox viruses; 2 of these, a putative endothelin-like protein and a MHC class I-like protein are likely to have immunomodulatory roles.

Keywords Poxvirus · Next-generation sequencing · NGS · Batpox · Eptesipoxvirus

Introduction

Poxviruses have dsDNA genomes that range from 130 to 360 kb and are sufficiently large that they can be seen by light microscopy. The *Poxviridae* family is divided into the *Entomopoxvirinae* subfamily of viruses that infect insects and the *Chordopoxvirinae* subfamily of viruses that infect vertebrates. According to the International Committee on Taxonomy of Viruses (ICTV) 2016 Release [1], 11 genera have been created to classify the Chordopoxviruses, but several viruses that are unclassified are likely to require new genera. Many of these genera contain only a few virus species, the exceptions are the *Orthopoxvirus* and *Avipoxvirus* genera. Orthopoxviruses have been extensively studied as models of poxvirus biology and the host immunological response because this group includes variola virus and vaccinia virus that are the agent of smallpox and the virus used as the smallpox vaccine, respectively. In contrast, avipoxvirus infections that have been recognized in a large number of wild and domestic birds have been widely surveyed by PCR techniques, but there is very limited sequence data available for these isolates.

As many of the individual virus names suggest, poxviruses infect a large array of hosts. Many poxviruses have been named after the animal from which they were first isolated; however, these animals may not be the natural reservoir hosts for these pathogens. For example, there are 3 groups of orthopoxviruses that have been named cowpox,

Edited by Joachim Jakob Bugert.

Electronic supplementary material The online version of this article (doi:10.1007/s11262-017-1485-4) contains supplementary material, which is available to authorized users.

✉ Chris Upton
cupton@uvic.ca

[1] Department of Biochemistry and Microbiology, University of Victoria, Victoria, BC, Canada

[2] The National Center for Emerging and Zoonotic Infectious Diseases Centers for Disease Control and Prevention, Atlanta, GA, USA

but these appear to be distinct species and are thought to be typically spread by rodents and possibly other small mammals [2]. Others such as goatpox virus, sheeppox virus, and lumpy skin disease virus (cattle) do not have a strictly limited host range and have been used as vaccines in the different host species [3]. Thus, it is important to note that the natural hosts of bat-isolated poxviruses may not be bats. However, poxviruses found in bats are of special interest because bats are associated with a number of known zoonotic disease viruses, including the coronaviruses that cause SARS [4] and MERS [5], the rhabdovirus that causes rabies [6], the filoviruses ebola [7], and Marburg viruses [8], and Australian bat lyssaviruses that cause rabies-like symptoms [9]. In 2013, 3 reports described the association of poxviruses with bats on 3 continents. *Eptesipoxvirus* (EPTV) was isolated in North America [10], *Eidolon helvum poxvirus 1* in West Africa [11], and a third virus was identified only by electron microscopy in South Australia [12]. Since only small amounts of sequence data were obtained from the first 2 viruses, the first complete genome of a bat-associated poxvirus was not published until 2016, from a Pteropox virus isolated in North Western Australia [13].

Bats, which have been categorized into approximately 1200 species, form the second most diverse animal taxon (*Chiroptera* order) after rodents. So far, bat-isolated poxviruses have been isolated from both the megabat and microbat suborders. These suborders differ significantly; megabats are large fruit bats with big eyes, small ears (always lack tails) that are found in tropical regions, whereas the small microbats have the opposite features, eat insects, and use echolocation [14, 15]. Thus, bats comprise a huge and diverse mammalian reservoir for poxviruses that may pose a risk to humans through zoonotic transmission [16]. Additionally, bat poxviruses may provide a genetic reservoir that, through recombination with other mammalian poxviruses, could generate new human pathogens [17]. It has been hypothesized that recombination events may have been involved in the development of variola virus as a human-specific pathogen [18]. Here, we aim to expand the available genome data and advance the study of bat-associated poxviruses. This paper presents the complete genomic sequence and annotation of EPTV, and establishes its potential relationship to other recognized poxvirus genera.

Materials and methods

Virus isolation and DNA preparation

The virus was isolated from a big brown bat (*Eptesicus fuscus*) that was treated at a wildlife rehabilitation center in the state of Washington due to joint swellings which impeded its ability to fly. Samples from an elbow joint were homogenized in sterile phosphate-buffered saline using a tissue grinder. Green monkey kidney epithelial cells (BSC40) were infected with 10 microliters of the homogenate and harvested after 95% of the cells were infected [10]. Viral nucleic acid was extracted from the harvested material using the Qiagen EZ1 Advanced XL. The sequencing library was prepared for use with the Illumina platform (HiSeq 2500; www.illumina.com/).

Sequence quality control and assembly

The single-ended raw Illumina EPTV sequence read file was subjected to quality control using Taxonomer, a metagenomics tool that assigns sequencing reads to taxonomic categories (taxonomer.iobio.io) [19]. At least 40% of the sequences were determined to be non-viral contamination (human, bacterial, and unknown sequences). Initially, the raw sequence reads that mapped to human genome (build 38) using the Burrows–Wheeler Alignment (BWA) tool [20] were removed, and the remaining reads were extracted using SAMtools and Fastqutils [21] prior to assembly by MIRA [22]. Second, using the same protocol, a separate assembly trial was performed with *Eptesicus fuscus* (bat host) sequences removed. However, both trials failed to generate a complete poxvirus genome contig, probably due to excessive read removals. Therefore, the final assembly was performed directly with the original raw read file with no decontamination process using designated "genome, denovo, accurate" parameters in MIRA. This contig was mapped to the original read file using the Tanoti short read aligner (http://www.bioinformatics.cvr.ac.uk/tanoti.php), and subsequently visualized in Tablet [23] for coverage analyses and manual base-calling.

Viral Bioinformatics Resource Centre

Genome annotation and analysis was performed with tools from the bioinformatics suite developed at the Viral Bioinformatics Resource Centre [24] (http://virology.uvic.ca/). The Genome Annotation Transfer Utility (GATU) [25] uses a reference genome to automatically annotate EPTV genes with clear orthologs in the reference. Other possible genes are presented to the annotator for further characterization and to make the final decision on annotation. Generally, our approach to annotation was conservative with ORFs smaller than 50 codons or overlapping with known genes being ignored. EPTV genes were preferentially referenced to VACV-Cop, with DPV-W1170_84 and CPXV-BR used second and third when an ortholog group does not exist in the prior species, respectively.

Virus orthologous clusters (VOCs) [26] and Base-By-Base (BBB) [27] were used to create MSAs of DNA and proteins.

Phylogenetic analyses and poxviral sequences used

Sequences used were extracted from the following virus strains using the VOCs database: EPTV-WA (KY747497); PTPV-Aus (KU980965); BPSV-BV-AR02 (NC_005337); CNPV-VR111 (NC_005309); COTV-SPAn232 (HQ64 7181); CPXV-BR (NC_003663); CRV-ZWE (NC_0080 30); DPV-W1170_84 (AY689437); FWPV-Iowa (NC_002 188); GTPV-Pellor (NC_004003); MOCV-st1 (NC_001 731); MYXV-Lau (NC_001132); ORFV-SA00 (NC_0053 36); RFV-Kas (NC_001266); RCNV-Herman (KP143769); SQPV-Red_squirrel_UK (HE601899); SWPV-Neb (NC_003389); TANV-KEN (NC_009888); TKPV_HU1124 (NC_028238.1); VARV-GBR44_harv (DQ4414 44); YMTV-Amano (NC_005179); YLDV-Davis (NC_002642); YKV-DakArB_4268 (HQ849551).

Seven conserved poxviral proteins (previously used in [10] and [13]): RPO147, RAP94, mRNA capping enzyme large subunit, P4a precursor, RPO132, VETF-L, and DNA primase were extracted from VOCs and aligned using MUSCLE [28] in Base-By-Base. MEGA6.06 was used to create a Maximum-likelihood tree using the LG model with the G + I rate parameter [29].

Results

Genome assembly and gene annotation

The initial assembly of the unfiltered Illumina single read sequencing data by MIRA [22] generated 2 large poxvirus-specific contigs (121,151, and 43,742 nt), which overlapped by 54 nt. After manually joining the 2 contigs and correction/matching of the genome Inverted Terminal Repeats (ITRs), a draft genome of 176,491 nt containing 2 copies of 11,999 nt ITRs was obtained. The mapping of raw reads against this draft sequence gave an average coverage of 150x, which was consistent throughout the genome except for 1 region (134,421–134,514 nt). This region, which showed a dramatic increase in coverage of 411x, was found to consist of 2 tandem repeats of about 54 nt each. Since the Illumina reads, which were approximately 71 nt on average, were unable to resolve the true number of repeats in the region, this region was amplified by PCR and re-sequenced using the Sanger method [30]. This process determined that this region consisted of six repeats rather than 2 and created a final EPTV genome sequence of 176,688 nt.

A conservative approach was taken to genome annotation to avoid over-annotation of ORFs that are unlikely to represent functional genes. ORFs less than 50 codons or overlapping more than 25% with well-characterized genes were not considered for annotation unless supported by other evidence. A total of 191 genes were annotated (Table 1), of which 13 are associated with each of the ITRs. EPTV-013 and EPTV-179 span the ITR junctions, the former being a N-terminus truncated version of the EPTV-179 gene that is predicted to encode an ankyrin-like protein. The GenBank accession number for the genome sequence of EPTV-WA is KY747497.

Phylogeny

The relationship between EPTV and the other orthopoxviruses was determined by (1) using MUSCLE [28] to align the result of concatenation of 7 conserved core proteins, (2) manual editing of the multiple sequence alignment (MSA) in base-by-base [27], (3) generation of an unrooted phylogenetic tree by MEGA v6.06 using maximum-likelihood [29]. The resulting radiation tree illustrates that EPTV branches off the backbone independent of all other viruses (Fig. 1). The genetic distance supports that EPTV should be placed into a new genus, which was previously proposed to be called *Chiropoxvirus* [10]. Thus, the 3 poxviruses isolated from bats (EPTV, PTPV, EHPV1) on different continents, each belong to a different poxvirus genus. Although bootstrap analysis strongly supports the position of EPTV in this tree, this is at odds with an earlier paper [10] which had indicated that EPTV was more closely related to Cotia virus (COTV) and that the pair branched off the backbone to form their own clade. When we analyzed the concatenated protein sequences used by Emerson et al. [10], we found that the protein sequences of the P4a precursor were shuffled, and placed with the concatenated sequences of the wrong virus species. Importantly, this switched the COTV and EPTV proteins with orthologous sequences from ORFV and BPSV, respectively, and this was sufficient, when the tree was generated, to create the impression that COTV and EPTV were within a single clade due to the close relationship of the parapoxviruses. Since the inclusion of the GC-rich viruses in the MSA results in the generation of a very large number of gaps and these frequently introduce mis-alignment around the indels, we also rebuilt the MSA omitting these sequences and recalculated the tree. The relationship between EPTV and neighboring viruses was identical to that shown in Fig. 1.

EPTV branches off the common backbone of the phylogenetic tree between the viruses that have been referred to as “clade II” viruses and the orthopoxviruses. The term clade II was coined to enable distinction between the

Table 1 EPTV genome annotation

Similar to clade II
Clade II-specific orthologs
Clade II orthologs not in orthopox

Similar to orthopoxvirus
Orthopox-specific orthologs
Orthopox orthologs not in clade II

Potential HGT genes
Novel EPTV genes
Conserved in Chordopox

Gene #	ORF position	AA #	Putative gene function	Orthologs
EPTV-001	620-141	159	Hypothetical protein	VACV-Cop-B15R
EPTV-002	1694-732	320	Serpin 2/CrmA (host range)	VACV-Cop-B13R
EPTV-003	2410-1727	227	Hypothetical protein	DPV-84-009
EPTV-004	3472-2474	332	IL-1 receptor-like protein	DPV-84-015
EPTV-005	3989-3510	159	Hypothetical protein	DPV-84-159
EPTV-006	4934-4029	301	Tyrosine protein kinase-like protein	DPV-84-158
EPTV-007	5683-4991	230	ER-localized apoptosis regulator (host range)	VACV-Cop-B9R
EPTV-008	6245-5769	158	Hypothetical protein	Unique to EPTV
EPTV-009	6804-6325	159	Hypothetical protein	VACV-Cop-B15R
EPTV-010	8696-6963	577	Ankyrin repeat-containing protein (host range)	CPXV-BR-025
EPTV-011	11040-9211	609	Ankyrin repeat-containing protein	DPV-84-019
EPTV-012	11599-11090	169	IFN resistance, eIF2a-like PKR inhibitor (host range)	VACV-Cop-K3L
EPTV-013	11992-11636	118	Ankyrin repeat protein fragment	DPV-84-019
EPTV-014	12864-12010	284	Monoglyceride lipase	VACV-Cop-K5L/K6L
EPTV-015	13183-12941	80	Secreted EGF-like protein	VACV-Cop-C11R
EPTV-016	13689-13189	166	Mitochondria anti-apoptotic factor (host range)	DPV-84-022
EPTV-017	14167-13742	141	dUTPase	VACV-Cop-F2L
EPTV-018	14598-14194	134	IFN-inducible protein (host range)	DPV-84-024
EPTV-019	15616-14642	324	Ribonucleotide reductase small subunit	VACV-Cop-F4L
EPTV-020	16786-15656	376	F5L membrane protein	VACV-Cop-F5L
EPTV-021	17026-16808	72	Hypothetical protein	Unique to EPTV
EPTV-022	17287-17063	74	Hypothetical protein	VACV-Cop-F7L
EPTV-023	17537-17346	63	Cytoplasmic protein	VACV-Cop-F8L
EPTV-024	18031-17645	128	Hypothetical protein	Unique to EPTV
EPTV-025	18714-18067	215	S-S bond formation pathway protein	VACV-Cop-F9L
EPTV-026	20017-18704	437	Ser/Thr protein kinase	VACV-Cop-F10L
EPTV-027	21337-20039	432	RhoA signalling inhibitor, virus release protein	VACV-Cop-F11L
EPTV-028	23302-21362	646	EEV maturation protein	VACV-Cop-F12L
EPTV-029	24456-23341	371	Palmitylated EEV membrane glycoprotein	VACV-Cop-F13L
EPTV-030	24703-24476	75	F14L conserved protein	VACV-Cop-F14L
EPTV-031	24950-24750	66	Hypothetical protein	Unique to EPTV
EPTV-032	25594-25148	148	F15L conserved protein	VACV-Cop-F15L
EPTV-033	26362-25682	226	Conserved non-functional serine recombinase	VACV-Cop-F16L
EPTV-034	26423-26737	104	DNA-binding phosphoprotein (VP11)	VACV-Cop-F17R
EPTV-035	28149-26731	472	Poly (A) polymerase catalytic subunit (VP55)	VACV-Cop-E1L
EPTV-036	30364-28166	732	IEV morphogenesis	VACV-Cop-E2L
EPTV-037	31138-30422	238	dsRNA-binding PKR inhibitor (host range)	VACV-Cop-E3L
EPTV-038	31903-31175	242	RNA polymerase subunit (RPO30)	VACV-Cop-E4L
EPTV-039	32072-33775	567	IMV protein, virion morphogenesis	VACV-Cop-E6R
EPTV-040	33799-34611	270	ER-localized membrane protein, virion core protein	VACV-Cop-E8R
EPTV-041	37628-34608	1006	DNA polymerase	VACV-Cop-E9L
EPTV-042	37661-37951	96	Sulfhydryl oxidase (FAD-linked)	VACV-Cop-E10R
EPTV-043	38365-37943	140	Virion core protein	VACV-Cop-E11L
EPTV-044	40424-38349	691	Virulence, modulates Raf/MEK/ERK pathway	VACV-Cop-O1L
EPTV-045	40795-40481	104	Nonessential glutaredoxin	VACV-Cop-O2L
EPTV-046	41842-40910	310	DNA-binding core protein	VACV-Cop-I1L
EPTV-047	42064-41843	73	IMV membrane protein	VACV-Cop-I2L
EPTV-048	42877-42065	270	ssDNA-binding phosphoprotein	VACV-Cop-I3L
EPTV-049	45218-42930	762	Ribonucleotide reductase large subunit	VACV-Cop-I4L
EPTV-050	45491-45255	78	IMV protein (VP13)	VACV-Cop-I5L
EPTV-051	46654-45509	381	Telomere-binding protein	VACV-Cop-I6L
EPTV-052	47933-46647	428	Viral core cysteine protease	VACV-Cop-I7L

Table 1 continued

EPTV-053	47939-49978	679	RNA-helicase, DExH-NPH-II	VACV-Cop-I8R
EPTV-054	51755-49965	596	Insulin metalloproteinase-like protein	VACV-Cop-G1L
EPTV-055	52084-51752	110	Entry/fusion complex component	VACV-Cop-G3L
EPTV-056	52078-52746	222	Late transcription elongation factor (VLTF)	VACV-Cop-G2R
EPTV-057	53090-52713	125	Thioredoxin-like protein	VACV-Cop-G4L
EPTV-058	53093-54409	438	FEN1-like nuclease	VACV-Cop-G5R
EPTV-059	54409-54600	63	RNA polymerase subunit (RPO7)	VACV-Cop-G5.5R
EPTV-060	54604-55143	179	NLPc/P60 superfamily protein	VACV-Cop-G6R
EPTV-061	56212-55109	367	Virion structural phosphoprotein, early morphogenesis	VACV-Cop-G7L
EPTV-062	56241-57023	260	Late transcription factor (VLTF-1)	VACV-Cop-G8R
EPTV-063	57039-58064	341	Myristylated entry/fusion protein	VACV-Cop-G9R
EPTV-064	58065-58814	249	Myristylated IMV envelope protein	VACV-Cop-L1R
EPTV-065	58847-59116	89	Crescent membrane/immature virion protein	VACV-Cop-L2R
EPTV-066	60073-59108	321	Internal virion protein	VACV-Cop-L3L
EPTV-067	60098-60856	252	DNA-binding virion protein (VP8)	VACV-Cop-L4R
EPTV-068	60871-61275	134	IMV protein, entry/fusion	VACV-Cop-L5R
EPTV-069	61217-61663	148	IMV membrane protein, virion morphogenesis	VACV-Cop-J1R
EPTV-070	61689-62219	176	Thymidine kinase	VACV-Cop-J2R
EPTV-071	62305-62916	203	Type I IFN inhibitor (host range)	VACV-Cop-C7L
EPTV-072	62991-63992	333	Poly (A) polymerase small subunit (VP39)	VACV-Cop-J3R
EPTV-073	63907-64464	185	RNA polymerase subunit (RPO22)	VACV-Cop-J4R
EPTV-074	64873-64466	135	IMV membrane protein, entry/fusion	VACV-Cop-J5L
EPTV-075	64979-68836	1285	RNA polymerase subunit (RPO147)	VACV-Cop-J6R
EPTV-076	69351-68833	172	Tyr/Ser kinase, virus assembly, IFN-gamma inhibitor	VACV-Cop-H1L
EPTV-077	69365-69937	190	Entry/fusion IMV protein	VACV-Cop-H2R
EPTV-078	70957-69944	337	IMV heparin-binding surface protein (p35)	VACV-Cop-H3L
EPTV-079	73348-70961	795	RNA polymerase-associated protein (RAP94)	VACV-Cop-H4L
EPTV-080	73557-74210	217	Late transcription factor (VLTF-4)	VACV-Cop-H5R
EPTV-081	74229-75170	313	DNA topoisomerase type I	VACV-Cop-H6R
EPTV-082	75202-75651	149	Crescent membrane/immature virion protein	VACV-Cop-H7R
EPTV-083	75694-78228	844	mRNA capping enzyme large subunit	VACV-Cop-D1R
EPTV-084	78633-78190	147	Virion core protein	VACV-Cop-D2L
EPTV-085	78632-79375	247	Virion core protein	VACV-Cop-D3R
EPTV-086	79372-80028	218	Uracil DNA glycosylase, DNA pol processivity factor	VACV-Cop-D4R
EPTV-087	80062-82425	787	NTPase, DNA primase	VACV-Cop-D5R
EPTV-088	82422-84329	635	Early transcription factor small subunit (VETF-s)	VACV-Cop-D6R
EPTV-089	84363-84875	170	RNA polymerase subunit (RPO18)	VACV-Cop-D7R
EPTV-090	85682-84810	290	Carbonic anhydrase, GAG-binding MV membrane protein	VACV-Cop-D8L
EPTV-091	85740-86408	222	mRNA decapping enzyme	VACV-Cop-D9R
EPTV-092	86386-87171	261	mRNA decapping enzyme	VACV-Cop-D10R
EPTV-093	89052-87145	635	ATPase, NPH1	VACV-Cop-D11L
EPTV-094	89948-89085	287	mRNA capping enzyme small subunit	VACV-Cop-D12L
EPTV-095	91661-89982	559	Trimeric virion coat protein (rifampicin resistance)	VACV-Cop-D13L
EPTV-096	92142-91687	151	Late transcription factor (VLTF-2)	VACV-Cop-A1L
EPTV-097	92847-92173	224	Late transcription factor (VLTF-3)	VACV-Cop-A2L
EPTV-098	93074-92844	76	S-S bond formation pathway protein	VACV-Cop-A2.5L
EPTV-099	95085-93094	663	P4b precursor	VACV-Cop-A3L
EPTV-100	95713-95141	190	Putative membrane-associated virion core protein (p39)	VACV-Cop-A4L
EPTV-101	95751-96260	169	RNA polymerase subunit (RPO19)	VACV-Cop-A5R
EPTV-102	97375-96257	372	Virion morphogenesis core protein	VACV-Cop-A6L
EPTV-103	99543-97399	714	Early transcription factor large subunit (VETF-L)	VACV-Cop-A7L
EPTV-104	99610-100485	291	Intermediate transcription factor (VITF-3s)	VACV-Cop-A8R
EPTV-105	100724-100482	80	IMV membrane protein, early morphogenesis	VACV-Cop-A9L
EPTV-106	103451-100725	908	P4a precursor	VACV-Cop-A10L
EPTV-107	103466-104401	311	Viral membrane formation	VACV-Cop-A11R
EPTV-108	104910-104398	170	Virion core and cleavage processing protein	VACV-Cop-A12L
EPTV-109	105224-105003	73	IMV membrane protein, virion maturation	VACV-Cop-A13L
EPTV-110	105572-105291	93	IMV membrane protein, essential	VACV-Cop-A14L

Table 1 continued

EPTV-111	105750-105589	53	IMV membrane protein, non-essential	VACV-Cop-A14.5L
EPTV-112	106033-105740	97	Core protein	VACV-Cop-A15L
EPTV-113	107159-106017	380	Myristylated protein, essential for entry/fusion	VACV-Cop-A16L
EPTV-114	107747-107160	195	IMV membrane protein	VACV-Cop-A17L
EPTV-115	107762-109201	479	DNA helicase, transcript release factor	VACV-Cop-A18R
EPTV-116	109409-109191	72	Zn finger-like protein, late morphogenesis	VACV-Cop-A19L
EPTV-117	109754-109410	114	IMV membrane protein, entry/fusion	VACV-Cop-A21L
EPTV-118	109753-111030	425	DNA polymerase processivity factor	VACV-Cop-A20R
EPTV-119	111014-111565	183	Holliday junction resolvase	VACV-Cop-A22R
EPTV-120	111562-112722	386	Intermediate transcription factor (VITF-3L)	VACV-Cop-A23R
EPTV-121	112719-116219	1166	RNA polymerase subunit (RPO132)	VACV-Cop-A24R
EPTV-122	119230-116222	1002	A-type inclusion protein	VACV-Cop-A25L
EPTV-123	121205-119274	643	P4c precursor	VACV-Cop-A26L
EPTV-124	121605-121261	114	IMV membrane protein, fusion	VACV-Cop-A27L
EPTV-125	122022-121606	138	IMV membrane protein, entry	VACV-Cop-A28L
EPTV-126	122938-122036	300	RNA polymerase subunit (RPO35)	VACV-Cop-A29L
EPTV-127	123149-122922	75	IMV protein	VACV-Cop-A30L
EPTV-128	123355-123855	166	Hypothetical protein	VACV-Cop-A31R
EPTV-129	124635-123871	254	ATPase/DNA packaging protein	VACV-Cop-A32L
EPTV-130	124771-125334	187	C-type lectin-like EEV membrane phosphoglycoprotein	VACV-Cop-A33R
EPTV-131	125386-125886	166	C-type lectin like IEV/EEV membrane glycoprotein	VACV-Cop-A34R
EPTV-132	125923-126441	172	MHC class II antigen presentation inhibitor	VACV-Cop-A35R
EPTV-133	126481-127326	281	Concanavalin-like precursor	DPV-84-135
EPTV-134	127368-128099	243	EEV glycoprotein	DPV-84-136
EPTV-135	128142-128972	276	Hypothetical protein	VACV-Cop-A37R
EPTV-136	128996-129259	87	Hypothetical protein	Unique to EPTV
EPTV-137	129840-129256	194	Truncated CD47-like protein, integral membrane protein	VACV-Cop-A38L
EPTV-138	129868-130266	132	Myristylated protein	VACV-Cop-E7R
EPTV-139	130325-131077	250	Hypothetical protein	SQPV-1300
EPTV-140	131927-131070	285	Chemokine-binding protein	VACV-Cop-A41L
EPTV-141	132054-132455	133	Profilin-like protein, ATI-localized	VACV-Cop-A42R
EPTV-142	132843-132457	128	Hypothetical protein	Unique to EPTV
EPTV-143	132917-133141	74	Hypothetical protein	Unique to EPTV
EPTV-144	133348-134403	351	3 β-hydroxysteroid dehydrogenase/δ 5→4 isomerase	VACV-Cop-A44L
EPTV-145	134449-134673	74	Hypothetical protein	Unique to EPTV
EPTV-146	135385-134732	217	Immunoprevalent protein	VACV-Cop-A47L
EPTV-147	135466-136053	195	Thymidylate kinase	VACV-Cop-A48R
EPTV-148	136085-137770	561	DNA ligase-like protein	VACV-Cop-A50R
EPTV-149	137817-139394	525	BTB kelch-domain protein	VACV-Cop-A55R
EPTV-150	139422-140048	208	A52R-like family protein	DPV-84-148
EPTV-151	140099-140596	165	Hypothetical protein	Unique to EPTV
EPTV-152	140648-141679	343	Hypothetical protein	VACV-Cop-A51R
EPTV-153	141743-142393	216	Toll/IL-1 receptor-like protein, NFkB signaling inhibitor	VACV-Cop-A52R
EPTV-154	142985-142518	155	Hypothetical protein	Unique to EPTV
EPTV-155	143074-144687	537	BTB kelch-domain protein	VACV-Cop-A55R
EPTV-156	144715-145509	264	Hemagglutinin	VACV-Cop-A56R
EPTV-157	145529-146464	311	Ser/Thr protein kinase	VACV-Cop-B1R
EPTV-158	146495-147493	332	IL-1 receptor antagonist	VACV-Cop-C10L
EPTV-159	147510-148409	299	KilA-N/RING finger protein (host range)	CPXV-BR-023
EPTV-160	148441-149031	196	Partial schlafen-like protein	VACV-Cop-B2R
EPTV-161	149126-149839	237	EEV type-1 membrane glycoprotein (host range)	VACV-Cop-B5R
EPTV-162	149954-150385	143	Anti-apoptotic Bcl-2-like protein	VACV-Cop-N1L
EPTV-163	150436-151284	282	dsRNA binding PKR inhibitor	VACV-Cop-E3L
EPTV-164	151349-152353	334	Serpin 1 (host range)	VACV-Cop-C12L
EPTV-165	152461-152919	152	Hypothetical protein	VACV-Cop-B15R
EPTV-166	152947-153822	291	Tyrosine protein kinase-like protein	DPV-84-158
EPTV-167	153877-154884	335	IL-1 beta-receptor	VACV-Cop-B16R
EPTV-168	154911-156857	648	Ankyrin repeat protein	DPV-84-019

Table 1 continued

EPTV-169	156942-157568	208	Ankyrin repeat protein	DPV-84-014
EPTV-170	157626-158333	235	Alpha-amanitin target protein	VACV-Cop-N2L
EPTV-171	158383-159060	225	NFkB inhibitor	VACV-Cop-M2L
EPTV-172	159108-159341	77	Endothelin precursor	DPV-84-006
EPTV-173	159392-160078	228	NFkB inhibitor	VACV-Cop-M2L
EPTV-174	160085-160876	263	Secreted complement binding protein (host range)	VACV-Cop-C3L
EPTV-175	160944-161372	142	IL-18 binding protein	DPV-84-021
EPTV-176	161449-161637	62	Hypothetical protein	Unique to EPTV
EPTV-177	161603-162202	199	Truncated TNFa-receptor (host range)	VACV-Cop-B28R
EPTV-178	162245-163201	318	MHC class I-like protein	SQPV-0040
EPTV-179	163239-165053	604	Ankyrin repeat protein	DPV-84-019
EPTV-180	165090-165599	169	IFN resistance, eIF2a-like PKR inhibitor (host range)	VACV-Cop-K3L
EPTV-181	165649-167478	609	Ankyrin repeat protein	DPV-84-019
EPTV-182	167993-169726	577	Ankyrin repeat protein (host range)	CPXV-BR-025
EPTV-183	169885-170364	159	Hypothetical protein	VACV-Cop-B15R
EPTV-184	170444-170920	158	Hypothetical protein	Unique to EPTV
EPTV-185	171006-171698	230	ER-localized apoptosis regulator (host range)	VACV-Cop-B9R
EPTV-186	171755-172660	301	Tyrosine protein kinase-like protein	DPV-84-158
EPTV-187	172700-173179	159	Hypothetical protein	DPV-84-159
EPTV-188	173217-174215	332	IL-1 receptor-like protein	DPV-84-015
EPTV-189	174279-174962	227	Hypothetical protein	DPV-84-009
EPTV-190	174995-175957	320	Serpin 2/CrmA (host range)	VACV-Cop-B13R
EPTV-191	176069-176548	159	Hypothetical protein	VACV-Cop-B15R

Ortholog families are represented by VACV-Cop designations; where the ortholog is lacking in VACV-Cop, the DPV-84 and CPXV-BR gene numbers are used, respectively. Inverted Terminal Repeat regions are bolded

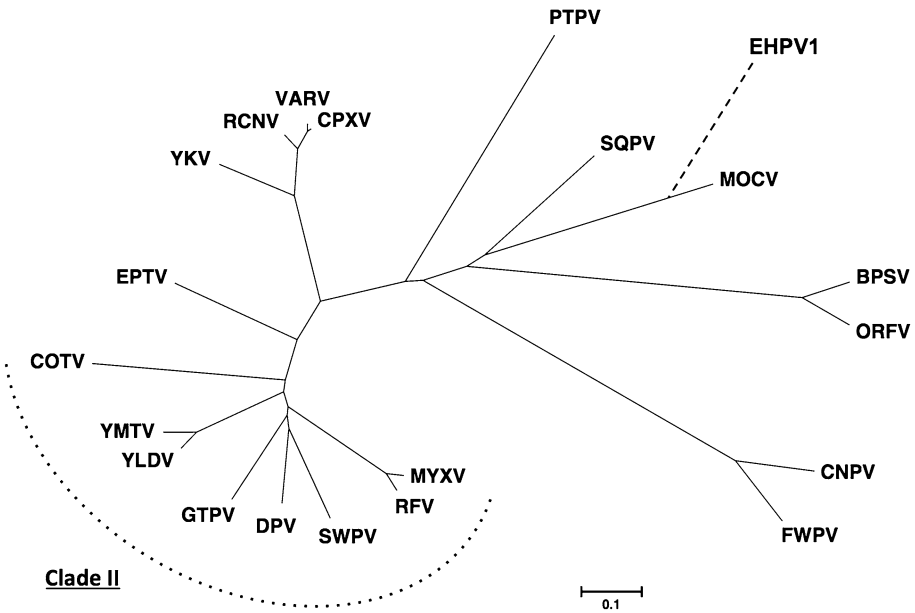

Fig. 1 Maximum-likelihood phylogenetic tree of EPTV with representatives of the *Chordopoxvirinae* subfamily; the MSA consisted of concatenated sequences of 7 conserved proteins: RPO147, RAP94, mRNA capping enzyme (large subunit), P4a precursor, RPO132, VETF-L, and DNA primase

monophyly of the "clade II" capri-, sui-, cervid-, yata-, leporipoxviruses, and COTV, to that of the ancient clade containing the orthopoxviruses [31]; clade II has also been alternatively referred to as the CSYLC clade [32]. Calculation of the average % amino acid (aa) identity from the MSA used to generate Fig. 1 reveals that EPTV is slightly more similar to clade II viruses (Fig. 1; 73–76%) than orthopoxviruses (Fig. 1; 71–72%); the exception is COTV, which has a very long branch. This phylogenetic relationship was also supported by comparison of the EPTV proteins to all other poxvirus proteins using BLASTP [33]; the "best match" was DPV in 25% of the searches. Other "best matches" were usually clade II viruses, but approximately 20% of searches yielded an orthopoxvirus as the "best

match." This type of comparison is somewhat over-simplified because it does not consider whether genes actually exist in the various genera; however, the trend prompted us to use DPV gene numbers as a reference in the absence of a VACV-Cop ortholog (Table 1).

The comparison of EPTV proteins and gene sequences to their orthologs in other poxviruses reveals the importance of understanding exactly what data are being compared when evaluating relationships between viruses. Although the essential genes present in the various poxviruses are likely to have evolved along identical paths and give similar phylogenetic trees, examination of DNA and protein sequences of different genes can produce considerably different views as to the divergence between viruses. As an illustration, Fig. 2 shows the aa and nt % identity between 6 EPTV and VACV-Cop orthologous gene pairs. The aa identity ranges from 30 to 83%, whereas the nt identity ranges from 68 to 75%. These results show (1) different selection pressures are acting on different genes within a genome, (2) the huge variation in aa identity observed between different orthologs from the same pair of viruses, and (3) the compression in the difference values when comparing DNA sequences rather than protein sequences due to the 4 and 20 letter codes of DNA and proteins, respectively. This illustrates why phylogenetic trees are usually constructed from conserved essential genes: using DNA sequences if the organisms are closely related, and with aa sequences if they are more divergent.

Unique EPTV genes

Given the >75% aa identity between some EPTV proteins and their orthologs, it would appear straightforward to determine which genes are unique to EPTV; however, as noted above and in Fig. 2, there is considerable variation between the similarity of the different ortholog pairs. Therefore, instead of choosing an arbitrary cut-off for assigning orthologs, we evaluated % identity, % similarity (allowing chemically similar aa to be matched), protein motifs, secondary structure prediction, synteny, and gene size. Hence, 21 EPTV proteins were predicted to have poxviral orthologs despite having <30% aa identity (Online Resource 1). Further, eleven EPTV genes: -008 (ITR: -184), -021, -024, -031, -136, -142, -143, -145, -151, -154, -176 could not be matched confidently with any poxvirus counterpart. However, several of these genes are located in positions in the genome where orthologs are missing when compared to the arrangement of genes in VACV-Cop (Table 1). Although it is possible that the EPTV orthologs have diverged to such an extent that no recognizable sequence patterns are discernable, it is also possible that genes have been replaced in EPTV. Interestingly, none of this set of EPTV sequences have any significant similarity to any of the sequences in the bat genome database [34], as well as the current databases when analyzed by the BLAST suite of programs [33]. A further question concerning the annotation of these ORFs is whether they are likely to be functional genes. The size of these unique ORFs ranges from 62 to 165 codons, five being in the set of twelve smallest ORFs annotated in EPTV (Table 1). This, together with the fact that several of these predicted proteins have isoelectric points in the extreme tails (low and high) of the distribution observed for known poxvirus proteins (not shown), suggests that some of these annotated ORFs may not represent functional genes [35]. Others may provide EPTV with functions specific to promotion of virus replication within its host.

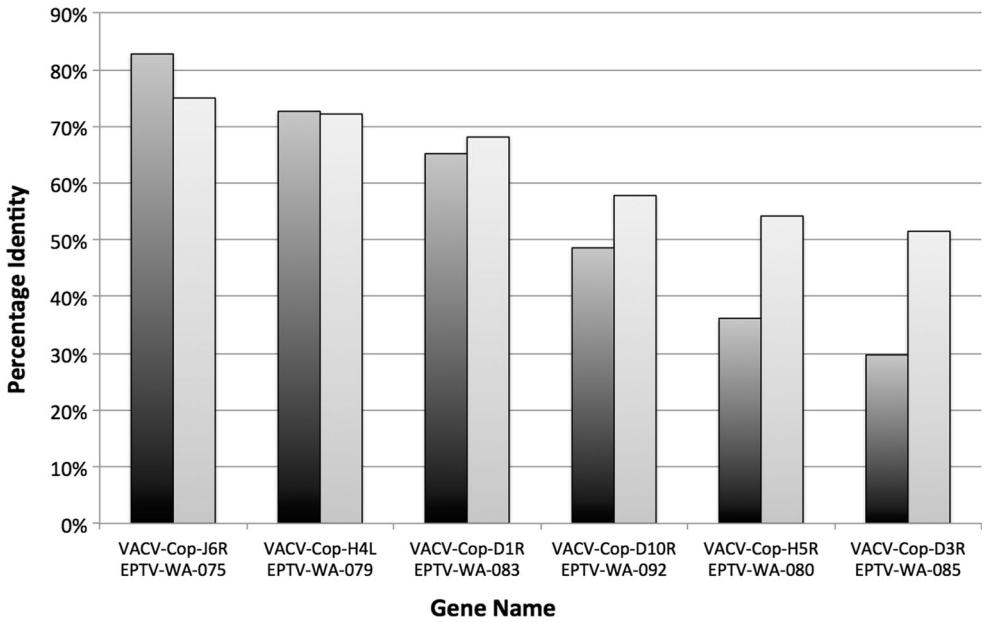

Fig. 2 Amino acid (*dark gray*) and nucleotide (*light gray*) percent identities between EPTV and VACV-Cop gene sets

Relationship with clade II poxviruses

As noted above, the % aa identity from the MSA of seven core proteins indicates that EPTV is slightly more similar to the clade II poxviruses. As shown in Table 1, EPTV possesses 7 genes that appear to only have clade II orthologs; however, some of these genes have one or more paralogs with different distributions among the various viruses that complicate the assignment of divergent genes to particular ortholog families. The EPTV genes -004 (-188; ITR), -018, -150, and -169 are predicted to have functions associated with virulence/host range (Table 1). EPTV genes -003 (-189; ITR), -005 (-187; ITR) have no known function, but are conserved in DPV and are 227 and 159 codons long, respectively; this suggests that they are likely functional genes in EPTV. EPTV-134 is unusual in that it is predicted to encode a clade II-specific structural protein that spans the outer membrane of intracellular enveloped virus (IEV) particles [36]. This clade II-specific IEV protein is located at the same position as orthopoxvirus ortholog A36R, which also encodes an IEV protein. Experimental evidence shows these clade II IEV orthologs (DPV-84-136) as functional orthologs of A36R, despite very low sequence identity, also induce actin tail formation, albeit via a different mechanism [37]. Using ScanSite3 [38], EPTV-134 is predicted to have five Nck-binding tyrosine motifs and 1 Grb2 binding motif; thus, it has the same signatures as DPV-84-136 [37] and is therefore likely to represent an ortholog that is otherwise clade II specific.

Further support for the closer association of EPTV with clade II viruses is that EPTV orthologs of VACV-Cop-C7L (-071) and -E7R (-138) genes (Table 1) are in positions that are syntenic only with their positions in clade II viruses [32].

Relationship with orthopoxviruses

EPTV also possesses several genes that are otherwise orthopoxvirus specific. EPTV-010 (ITR: -182) is an ortholog of CPXV-BR-025, which encodes an ankyrin-like protein (577 aa). It is curious that the neighboring EPTV-011 gene also encodes an ankyrin-like protein (609 aa), which is an ortholog of DPV-84-019 that is associated with clade II, avipox, and parapox viruses. Despite the physical proximity of these 2 genes, the aligned predicted protein sequences are only 23% identical. In addition, EPTV-010 and EPTV-011 have a different number of ankyrin repeats, 3 and 9, respectively. Although it is assumed that gene duplication events have created families of paralogous genes in poxviruses, at this level of sequence diversity it is very difficult to differentiate this from the possibility that some may be the result of separate gene capture events. COTV is the only other clade II chordopoxvirus (to date) that is believed to have a mixture of ankyrin orthologs from both orthopoxviruses and clade II viruses [26]. Together with COTV, EPTV also encodes a VACV-C3L homolog (secreted complement binding protein) that is otherwise exclusive to the orthopoxviruses.

Additional EPTV genes that are otherwise orthopoxvirus specific are scattered throughout the genome: EPTV-045, -141, -146, and -167 are orthologs of VACV-O2L, -A42R, -A47L, and -B16R, and encode a glutaredoxin-like protein, a profilin-like protein, the immunoprevalent protein, and an IL-1 beta-receptor-like protein, respectively. EPTV also encodes a chemokine-binding protein (EPTV-140) as well as a partial schlafen-like protein (EPTV-160) that have only been found in orthopoxviruses, yokapoxvirus (YKV), and the recently sequenced pteropoxvirus [13, 39].

Variably present genes

Above, we tried to categorize EPTV genes by which other viruses have orthologs; however, when the presence of a particular gene is inconsistent through different genera, it is also instructive to determine those genera that do not contain an ortholog. For example, the EPTV-122, -123, and -141 proteins that create or localize to an A-type inclusion body are absent from the clade II viruses. Similarly, EPTV encodes a ribonucleotide reductase large subunit (I4L) that is missing from all clade II viruses except swinepox (SWPV). In contrast, EPTV-006 (-185), -011 (-181), -016, -133, -166, -168, -175, and -179 are absent from the orthopoxviruses, but variably present in the clade II viruses and other chordopoxviruses (Table 1). This mosaic nature of EPTV genes makes the elucidation of evolutionary events interesting but more difficult.

A link between diverged F5L ortholog families

Although EPTV-019 (ribonucleotide reductase, small subunit) and EPTV-025 (S–S bond formation pathway protein) are clear orthologs of VACV-F4L and -F9L, the genes between these pairs have varied relationships. This is not unexpected because the region between F4L and F9L is also highly diverse between chordopoxviruses species, and genes here retain little or no identity with a VACV reference genome. EPTV-020 shows similarity with VACV-F5L (membrane protein) and MOCV/SQPV-003 throughout the sequence (26–29% aa ID with functional aa conserved), and peaks in identity with clade II DPV-84-027 ortholog group at the C-terminus. However, the 3 groups of protein showed no discernible similarity with each other prior to the discovery of EPTV-020 sequence (elaborated below). EPTV-021 (hypothetical gene), which could

encode a 72 aa polypeptide, is unique to EPTV and replaces VACV-F6L, that is itself a small hypothetical gene. Although EPTV-022 and -023 are predicted to encode small proteins of unknown function, they have some similarity to the gene products of VACV-F7L and F8L. EPTV-024, which is capable of encoding a 128 aa polypeptide, is another gene that is unique to EPTV.

Discovery of EPTV-020 revealed MOCV003L/SQPV003 and clade II DPV-84-027 ortholog family as divergent F5L proteins. Previously, the MOCV/SQPV orthologs had only 17% aa ID with VACV-F5L, while the clade II orthologs (DPV-84-027) that replaced the F5L position in the genome appeared to lack detectable sequence similarity to the established orthopox F5L orthologs. However, they actually share up to 39.5% aa identity with EPTV-020 at the C-terminal region, rather than just 22% aa identity with orthopoxviruses. This suggests that these F5L-positioned clade II orthologs (DPV-84-027) are likely N-terminus truncated versions of F5L orthologs with subsequent divergence. This is significant because it demonstrates a closer relationship between EPTV and clade II that is not visible in the phylogenetic tree or the 1% difference in identity observed in the MSA. Furthermore, this relationship, uncovered by analysis of the EPTV-020 sequence, supports the idea that the clade II gene at the F5L position evolved from the same ancestral F5L rather than being acquired by an independent gene capture event.

Other genes of interest

EPTV-172 encodes an endothelin-like polypeptide; only DPV has a similar gene. Eukaryotic endothelin is part of a multi-gene family that, along with at least 4 receptors, acts as vaso-constrictors [40]. However, as noted by others [41] there is currently no experimental evidence as to whether this viral ortholog acts in a manner similar to the endothelins or as an antagonist. Blocking endothelin function by binding a receptor without initiating signaling might be favorable to the virus due to a reduction of inflammation [42].

EPTV-139 is predicted to encode a cysteine-rich protein that has similarity to only 1 other poxvirus protein, SQPV-130. It is curious to see an AT-rich poxvirus, like EPTV, share a gene exclusively with a GC-rich virus like SQPV. SQPV-130 is a hypothetical protein of unknown function that was previously erroneously annotated as containing a death effector domain (Dr. C. McInnes, personal communication). Although these poxvirus proteins have only 23% aa identity, 10 cysteines are conserved over the 250 aa alignment despite the fact that these EPTV and SQPV genes have an A+T nucleotide composition of 82 and 30%, respectively. The proteins are not predicted to have N-terminal signal sequences, but since poxviruses encode proteins that create S–S bonds within the cytoplasm [43], these proteins may still be structurally dependent on these cysteines.

Although aa identity is low (25%), the EPTV-178 protein is clearly a member of the MHC class I family by virtue of the conservation of a characteristic pattern of cysteine residues. Several other poxviruses have similar genes, but it appears that at least some have been independently acquired, since, despite low identity scores, EPTV-178 and SQPV-004 proteins score best in BLASTP searches against microbat and squirrel MHC class I proteins, respectively. COTV also has an independently acquired MHC class I protein most similar to a homolog in a Southern/Central American wild cat [32].

Discussion

The reconstructed phylogeny shows that, unlike what was previously suggested, EPTV does not form a sister clade with COTV. Instead, it branches off the common backbone from the rest of the A+T % rich chordopoxviruses and forms its own genus between clade II poxviruses and orthopoxviruses. This effectively blurs the distinction between orthopoxviruses and clade II viruses. The term "clade II" arose to distinguish a clade of viruses (capri-, sui-, cervid-, yata-, leporipoxviruses, and COTV) from the orthopoxviruses which diverged at their last common node [31]. However, given the phylogenetic position of EPTV and its mosaic collection of orthologous clade II and orthopoxvirus genes (this characteristic is also exhibited by COTV to some extent) [32], the term may have outlived its usefulness. EPTV along with the recently sequenced pteropoxvirus [13] demonstrate the value of sequencing further poxvirus genomes to increase the information in phylogenetic trees.

In terms of genome statistics, the 176,688 nt genome size for EPTV is considered medium length among the chordopoxviruses. It is predicted to encode 191 genes and has one of the larger ITRs sequenced (>10,000 nt). It should be noted that EPTV and COTV genomes are the most A+T % rich of the chordopoxvirus genomes (76.4%). The reduced synteny, including gene loss, at the right end of the EPTV genome compared to VACV-Cop reflects the fact that many VACV genes in this region may be non-essential. Three immunomodulatory genes suspected to be the result of horizontal gene transfer have been found in this region (EPTV-139, -172, and -178), including 1 predicted to encode a MHC class I protein (Table 1).

Further examples of genomic variability towards the terminal regions of the EPTV genome include the presence of 3 and 8 novel EPTV ORFs at the left and right ends, respectively. Although these "unique" EPTV sequences

may encode new virulence factors, their small sizes cast doubt as to whether they produce novel polypeptides. However, it is likely that the process of generating novel genes (not paralogs) in poxviruses begins small, for example, the creation of a promoter region adjacent to a novel ORF in new DNA acquired by a horizontal transfer event or in a fragmented pseudogene (non-functional). Such an initiating event could be followed by the gain of further protein domains [44, 45].

With regard to the potential for cross-species infections, EPTV possesses 11 out of 12 sets of host range genes previously reviewed [46, 47] (missing K1L, which is only found in orthopoxviruses). Of these 11, unique sequence extensions and truncations are found associated with EPTV orthologs of PKR inhibitors (K3L and E3L), KilA-N/RING domain protein (p28/N1R) involved in ubiquitin or apoptosis inhibition, and tumor necrosis factor receptor family 2 (B28R). It is unknown how these altered sequences may impact the ability of the encoded protein to avert different host immune responses. For example, SPPV with a deletion of the kelch-like protein (SPPV-019) restored 100% survival rate in infected sheep, but the same deletion failed to reduce virulence in a VACV model [48]. Therefore, the mere possession of host range proteins alone cannot serve as a marker or indicator for cross-species infection without considering the diversity of viral sequences and host immune systems, as well as any combinatorial effect of gene sets. However, we can foresee some beneficial role for EPTV in having either an increased virulence that promotes infections, or an attenuation that could contribute to the host serving as a reservoir of different viruses. The latter seems to align with the fact that majority of bat-borne viruses exhibit non-severe nor fatal symptoms in their reservoir hosts [49].

The discovery of multiple poxviruses in bats leads to inquiries about the possibility of coevolution between the host and virus [50]. Bats, whose species have diverged 50 million years ago, include diverse species in the suborders of "megabats" and "microbats," differing both on the behavioral and molecular levels [14, 15]. We suggest that bat poxviruses, like other poxviruses shown to date, do not usually co-evolve with their hosts [51]. First of all, there is no evidence yet to suggest the individual bats as the natural hosts of these poxviruses (EPTV, PTVP, and EHPV1); surveys have yet to be conducted on these geographically distinct bats that screen for the prevalence of a long-harboring poxvirus pool in the population. Secondly, given bats are indeed common hosts of these poxviruses, the phylogeny (Fig. 1) clearly shows EPTV and PTPV branch off into individual genera that are not in proximity to each other on the phylogenetic tree the way bats from different suborders usually group; EHPV, represented by a partial RPO18 amino acid sequence, branches off yet another separate genus close to molluscum contagiosum virus (MOCV). This is consistent with the poxviral phylogenetic analysis where capripox- and parapoxviruses do not cluster into the same branch despite common hosts in even-toed hoofed mammals [51].

It is notable that A+T-rich EPTV (76%) groups with other A+T-rich viruses, while PTPV (66% A+T) and EHPV (N/A) both group closer to G+C-rich viruses. In fact, PTPV has been noted to possess 3 genes, also in the same position, observed only in G+C-rich viruses. Although this is a small sample size, it is intriguing that PTPV and EHPV were isolated from megabat hosts (*Pteropus scapulatus* and *Eidolon helvum*), whereas EPTV was isolated from a microbat (*Eptesicus fuscus*). However, there is no evidence for a correlation between virus and host genome nucleotide composition.

Lastly, we found no evidence of similarities between the genes shared by the bat-borne EPTV and PTPV that suggest common virulence mechanisms in bat hosts. Consistent with this theme, the clinical symptoms of EPTV in *Eptesicus fuscus* manifest in the form of joint swelling and increased lethargy [10], whereas *Pteropus scapulatus* infected with PTPV presented lesions on wing membranes [13]; neither type of symptoms was directly linked to fatality. Thus, it appears that these "bat-isolated poxviruses" do not have any common genes based on their related hosts.

Acknowledgements We thank the many University of Victoria students that helped build the Viral Bioinformatics Resource Centre, and Chad Smithson for all the knowledge he has provided. CU received an award from the Natural Sciences and Engineering Research Council (NSERC) (Discovery Grant No. 04953) to support this work.

Disclaimer The findings and conclusions in this report are those of the authors and do not necessarily represent the views of the Centers for Disease Control and Prevention.

Author contributions The following conceived of, or designed the study (GLE, DSC, YL, CU), performed research (SLT, YN, JG, KW, NGR), analyzed data (SLT, YN, JG), and wrote the manuscript (SLT, GLE, CU, YN).

Compliance with ethical standards

Conflicts of interest The authors declare that they have no conflict of interest.

Ethics statement This article does not contain any studies with human participants or animals performed by any of the authors.

Informed consent No human subjects were involved in this study.

References

1. M.J. Adams, E.J. Lefkowitz, A.M.Q. King, B. Harrach, R.L. Harrison, N.J. Knowles, A.M. Kropinski, M. Krupovic, J.H.

Kuhn, A.R. Mushegian, M. Nibert, S. Sabanadzovic, H. Sanfaçon, S.G. Siddell, P. Simmonds, A. Varsani, F.M. Zerbini, A.E. Gorbalenya, A.J. Davison, Arch. Virol. **161**, 2921 (2016)
2. D.S. Carroll, G.L. Emerson, Y. Li, S. Sammons, V. Olson, M. Frace, Y. Nakazawa, C.P. Czerny, M. Tryland, J. Kolodziejek, N. Nowotny, M. Olsen-Rasmussen, M. Khristova, D. Govil, K. Karem, I.K. Damon, H. Meyer, PLoS ONE **6**, e23086 (2011)
3. E.S.M. Tuppurainen, C.R. Pearson, K. Bachanek-Bankowska, N.J. Knowles, S. Amareen, L. Frost, M.R. Henstock, C.E. Lamien, A. Diallo, P.P.C. Mertens, Antivir Res. **109**, 1 (2014)
4. X.-Y. Ge, J.-L. Li, X.-L. Yang, A.A. Chmura, G. Zhu, J.H. Epstein, J.K. Mazet, B. Hu, W. Zhang, C. Peng, Y.-J. Zhang, C.-M. Luo, B. Tan, N. Wang, Y. Zhu, G. Crameri, S.-Y. Zhang, L.-F. Wang, P. Daszak, Z.-L. Shi, Nature **503**, 535 (2013)
5. Z.A. Memish, N. Mishra, K.J. Olival, S.F. Fagbo, V. Kapoor, J.H. Epstein, R. AlHakeem, A. Durosinloun, M. Al Asmari, A. Islam, A. Kapoor, T. Briese, P. Daszak, A.A. Al Rabeeah, W.I. Lipkin, Emerg. Infect. Dis. **19**, 1819 (2013)
6. T. Binger, A. Annan, J.F. Drexler, M.A. Müller, R. Kallies, E. Adankwah, R. Wollny, A. Kopp, H. Heidemann, D. Dei, F.C. Agya-Yao, S. Junglen, T. Feldt, A. Kurth, S. Oppong, Y. Adu-Sarkodie, C. Drosten, J. Virol. **89**, 4588 (2015)
7. E.M. Leroy, A. Epelboin, V. Mondonge, X. Pourrut, J.-P. Gonzalez, J.-J. Muyembe-Tamfum, P. Formenty, Vector-Borne Zoonotic Dis. **9**, 723 (2009)
8. J.S. Towner, B.R. Amman, T.K. Sealy, S.A.R. Carroll, J.A. Comer, A. Kemp, R. Swanepoel, C.D. Paddock, S. Balinandi, M.L. Khristova, P.B.H. Formenty, C.G. Albarino, D.M. Miller, Z.D. Reed, J.T. Kayiwa, J.N. Mills, D.L. Cannon, P.W. Greer, E. Byaruhanga, E.C. Farnon, P. Atimnedi, S. Okware, E. Katongole-Mbidde, R. Downing, J.W. Tappero, S.R. Zaki, T.G. Ksiazek, S.T. Nichol, P.E. Rollin, PLoS Pathog. **5**, e1000536 (2009)
9. A.R. Gould, A.D. Hyatt, R. Lunt, J.A. Kattenbelt, S. Hengstberger, S.D. Blacksell, Virus Res. **54**, 165 (1998)
10. G.L. Emerson, R. Nordhausen, M.M. Garner, J.R. Huckabee, S. Johnson, R.D. Wohrle, W.B. Davidson, K. Wilkins, Y. Li, J.B. Doty, N.F. Gallardo-Romero, M.G. Metcalfe, K.L. Karem, I.K. Damon, D.S. Carroll, Emerg. Infect. Dis. **19**, 1002 (2013)
11. K.S. Baker, R.M. Leggett, N.H. Bexfield, M. Alston, G. Daly, S. Todd, M. Tachedjian, C.E.G. Holmes, S. Crameri, L.-F. Wang, J.L. Heeney, R. Suu-Ire, P. Kellam, A.A. Cunningham, J.L.N. Wood, M. Caccamo, P.R. Murcia, Virology **441**, 95 (2013)
12. D.J. McLelland, T. Reardon, S. Bourne, C. Dickason, A. Kessell, W. Boardman, *Outbreak of Skin Nodules Associated with Riouxgolvania Beveridgei (Nematoda: Muspiceida) in the Southern Bentwing Bat (Miniopterus Schreibersii Bassanii), South Australia* (Wildlife Disease Association, 2013), pp. 1009–1013
13. M.A. O'Dea, S.-L. Tu, S. Pang, T. De Ridder, B. Jackson, C. Upton, J. Gen. Virol. **97**, 2363 (2016)
14. M.S. Springer, E.C. Teeling, O. Madsen, M.J. Stanhope, W.W. de Jong, Proc. Natl. Acad. Sci. USA **98**, 6241 (2001)
15. E.C. Teeling, M.S. Springer, O. Madsen, P. Bates, S.J. O'Brien, W.J. Murphy, Science **307**, 580 (2005)
16. C.H.C. Ricardo Moratelli, Memórias Do Instituto Oswaldo Cruz **110**, 1 (2015)
17. C.R. Parrish, E.C. Holmes, D.M. Morens, E.C. Park, D.S. Burke, C.H. Calisher, C.A. Laughlin, L.J. Saif, P. Daszak, Microbiol. Mol. Biol. Rev. **72**, 457 (2008)
18. C. Smithson, A. Purdy, A.J. Verster, C. Upton, PLoS ONE **9**, e91520 (2014)
19. S. Flygare, K. Simmon, C. Miller, Y. Qiao, B. Kennedy, T. Di Sera, E.H. Graf, K.D. Tardif, A. Kapusta, S. Rynearson, C. Stockmann, K. Queen, S. Tong, K.V. Voelkerding, A. Blaschke, C.L. Byington, S. Jain, A. Pavia, K. Ampofo, K. Eilbeck, G. Marth, M. Yandell, R. Schlaberg, Genome Biol. **17**(1), 111 (2016)
20. H. Li, R. Durbin, Bioinformatics **25**, 1754 (2009)
21. M.R. Breese, Y. Liu, Bioinformatics **29**, 494 (2013)
22. B. Chevreux, T. Pfisterer, B. Drescher, A.J. Driesel, W.E.G. Müller, T. Wetter, S. Suhai, Genome Res. **14**, 1147 (2004)
23. I. Milne, M. Bayer, G. Stephen, L. Cardle, D. Marshall, Methods Mol. Biol. **1374**, 253 (2016)
24. M. Da Silva, C. Upton, Methods Mol. Biol. **890**, 233 (2012)
25. V. Tcherepanov, A. Ehlers, C. Upton, BMC Genom. **7**, 150 (2006)
26. A. Ehlers, J. Osborne, S. Slack, R.L. Roper, C. Upton, Bioinformatics **18**, 1544 (2002)
27. W. Hillary, S.-H. Lin, C. Upton, Microb. Inform. Exp. **1**, 2 (2011)
28. R.C. Edgar, BMC Bioinform. **5**, 113 (2004)
29. K. Tamura, G. Stecher, D. Peterson, A. Filipski, S. Kumar, Mol. Biol. Evol. **30**, 2725 (2013)
30. F. Sanger, S. Nicklen, A.R. Coulson, Proc. Natl. Acad. Sci. USA **74**, 5463 (1977)
31. A.L. Hughes, R. Friedman, Mol. Phylogenet. Evol. **35**, 186 (2005)
32. P.P. Afonso, P.M. Silva, L.C. Schnellrath, D.M. Jesus, J. Hu, Y. Yang, R. Renne, M. Attias, R.C. Condit, N. Moussatché, C.R. Damaso, J. Virol. **86**, JVI.07162 (2012)
33. S.F. Altschul, W. Gish, W. Miller, E.W. Myers, D.J. Lipman, J. Mol. Biol. **215**, 403 (1990)
34. G. Zhang, C. Cowled, Z. Shi, Z. Huang, K.A. Bishop-Lilly, X. Fang, J.W. Wynne, Z. Xiong, M.L. Baker, W. Zhao, M. Tachedjian, Y. Zhu, P. Zhou, X. Jiang, J. Ng, L. Yang, L. Wu, J. Xiao, Y. Feng, Y. Chen, X. Sun, Y. Zhang, G.A. Marsh, G. Crameri, C.C. Broder, K.G. Frey, L.-F. Wang, J. Wang, Science **339**, 456 (2013)
35. C. Upton, Virus Genes **20**, 159 (2000)
36. G.L. Smith, M. Law, Virus Res. **106**, 189 (2004)
37. M.P. Dodding, M. Way, Cell Host Microbe **6**, 536 (2009)
38. J.C. Obenauer, L.C. Cantley, M.B. Yaffe, Nucl. Acids Res. **31**, 3635 (2003)
39. G. Zhao, L. Droit, R.B. Tesh, V.L. Popov, N.S. Little, C. Upton, H.W. Virgin, D. Wang, J. Virol. **85**, 10230 (2011)
40. A. Goraca, Endocr. Regul. **36**, 161 (2002)
41. C.L. Afonso, G. Delhon, E.R. Tulman, Z. Lu, A. Zsak, V.M. Becerra, L. Zsak, G.F. Kutish, D.L. Rock, J. Virol. **79**, 966 (2005)
42. B. Hocher, A. Schwarz, K.A. Fagan, C. Thöne-Reineke, K. El-Hag, H. Kusserow, S. Elitok, C. Bauer, H.-H. Neumayer, D.M. Rodman, F. Theuring, Am. J. Respir. Cell Mol. Biol. **23**, 19 (2012)
43. H.-P. Su, D.Y.-W. Lin, D.N. Garboczi, J. Virol. **80**, 7706 (2006)
44. J. Jin, X. Xie, C. Chen, J.G. Park, C. Stark, D.A. James, M. Olhovsky, R. Linding, Y. Mao, T. Pawson, Sci Signal **2**, ra76 (2009)
45. A. Shukla, R. Hilgenfeld, Virus Genes **50**, 29 (2015)
46. S.L. Haller, C. Peng, G. McFadden, S. Rothenburg, Infect. Genet. Evol. **21**, 15 (2014)
47. K.A. Bratke, A. McLysaght, S. Rothenburg, Infect. Genet. Evol. **14**, 406 (2013)
48. C.A. Balinsky, G. Delhon, C.L. Afonso, G.R. Risatti, M.V. Borca, R.A. French, E.R. Tulman, S.J. Geary, D.L. Rock, J. Virol. **81**, 11392 (2007)
49. H.-J. Han, H.-L. Wen, C.-M. Zhou, F.-F. Chen, L.-M. Luo, J.-W. Liu, X.-J. Yu, Virus Res. **205**, 1 (2015)
50. P.R. Murcia, K.S. Baker, Viruses **7**, 5571–5586 (2015). **6**, 1564 (2014)
51. A.L. Hughes, S. Irausquin, R. Friedman, Infect. Genet. Evol. **10**, 50 (2010)